

**Amendment to the Claims**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1-45 (Canceled)

Please enter the following new claims:

46. (New) A method, embodied in a computer program, for automated extraction data from a molecular array having features arranged in a regular pattern, the method comprising:
- receiving a number of images of the molecular array, each produced by scanning the molecular array to determine intensities of data signals emanating from discrete positions on a surface of the molecular array;
  - estimating initial positions of selected marker features within an image of the molecular array;
  - calculating refined positions of the selected marker features within the image of the molecular array;
  - using the refined positions of the selected marker features to compute an initial coordinate system for locating features of the molecular array in the number of images of the molecular array;
  - using the initial coordinate system to locate positions of strong features within one or more images of the molecular array;
  - refining the positions of strong features within the one or more images of the molecular array by analyzing data signal intensity values in regions of the one or more images of the molecular array that contain the strong features;
  - using the refined positions of strong features in the one or more images of the molecular array to calculate a refined coordinate system to locate positions of weak features within the number of images of the molecular array;
  - using the refined positions of strong features in the one or more images of the

USSN 10/648,819

2

molecular array to calculate a refined coordinate system to locate positions of local background regions surrounding all strong and weak features within the number of images of the molecular array; and

extracting data from strong features, and their respective local background regions, within the number of images of the molecular array using the refined positions of strong features within the number of images of the molecular array and extracting data from weak features, and their respective local background regions, within the number of images of the molecular array using locations for the weak features calculated from the refined coordinate system.

47. (New) The method of claim 46 wherein data signals emanating from discrete positions on the surface of the molecular array include:

fluorescent emission from fluorophores incorporated into molecules bound to features of the molecular array;

radiation emitted by radioisotopes incorporated into molecules bound to features of the molecular array; and

light emission from chemoluminescent moieties incorporated into molecules bound to features of the molecular array.

48. (New) The method of claim 46 wherein each image of the number of images comprise an array of pixels, each pixel having a data signal intensity value.

49. (New) The method of claim 48 wherein the features of the molecular array are arranged in a rectilinear grid, wherein corner features are selected as marker features, and wherein estimating initial positions of selected marker features within an image of the molecular array further includes:

calculating row and column vectors by considering the values of pixels in rows and columns of the image;

determining a first and last peak in the row and column vectors; and

using pixel coordinates of the first and last peaks in the row vector to determine horizontal coordinates of the corner features and using pixel coordinates of the first and last peaks in the column vector to determine vertical coordinates of the corner features.

50. (New) The method of claim 49 wherein row and column vectors are calculated for various orientations of an image produced by scanning the molecular array, and the orientation at which the column and row vectors have sharpest and most distinct peaks is selected as the orientation for subsequent calculations.

51. (New) The method of claim 48 wherein calculating refined positions of the selected marker features within the image of the molecular array further includes:

- for each selected marker feature,
- creating a binary image of the selected marker feature; and
- using blob analysis to determine refined pixel coordinates for the feature.

52. (New) The method of claim 51 wherein the binary image is generated by:

- creating a histogram of intensity values within an inner region of interest that contains pixels corresponding to the feature;

- selecting an intensity value as a threshold value from the histogram based on a ratio of a number of pixels having an intensity value greater than the threshold value divided by a total number of pixels within the inner region of interest; and

- for each pixel in the binary image, setting the value of the pixel to 1 if the intensity value of a corresponding pixel in the image is greater than or equal to the threshold value, and setting the value of the pixel to 0 if the intensity value of the corresponding pixel in the image is less than the threshold value.

53. (New) The method of claim 51 wherein the binary image is generated by:

- determining a median intensity value and standard deviation of intensity values of pixels within an outer region of interest that contains the pixels corresponding to the feature;

selecting as a threshold value an intensity value that is a number of standard deviations greater than the median intensity value;

for each pixel in the binary image, setting the value of the pixel to 1 if the intensity value of a corresponding pixel in the image is greater than or equal to the threshold value, and setting the value of the pixel to 0 if the intensity value of the corresponding pixel in the image is less than the threshold value.

54. (New) The method of claim 48 wherein a strong feature is a feature for which a largest blob produced during blob analysis of the feature is greater than or equal to a minimal size and less than a maximum size and located within a threshold distance from a center position of the feature estimated from the initial coordinate system.

55. (New) The method of claim 54 wherein blob analysis is used to refine the positions of strong features and their respective local background regions within the one or more images of the molecular array by analyzing data signal values in regions of the one or more images of the molecular array that contain the strong features.

56. (New) The method of claim 48 wherein linear regression analysis is used to calculate a refined coordinate system to locate positions of weak features and their respective local background regions within the number of images of the molecular array.

57. (New) The method of claim 56 wherein the linear regression analysis is carried out using two matrix multiplication steps, in a first step calculating a coefficient matrix by multiplying together a transpose of a pseudo-inverse of a strong matrix containing indices of strong features with a centroid matrix containing coordinates of centroids of strong features, and in the second step calculating a fit positions matrix by: multiplying together a transposed feature indices matrix and the coefficient matrix calculated in the first step.

58. (New) The method of claim 46 wherein extracting data from strong features, and their respective local background regions, within the number of images of the molecular array using the refined positions of strong features within the number of images of the molecular array and extracting data from weak features, and their respective local background regions, within the number of images of the molecular array using locations for the weak features calculated from the refined coordinate system further includes:

- determining which pixels are included in the extraction of signal from each feature or local background region using statistical methods for pixel outlier identification;

- determining averages and variances of data signal intensities for features of the molecular array and covariances for one or more pairs of data signal intensities;

- determining averages and variances for background data signal intensities;

- determining background-subtracted averages and variances of data signal intensities for features of the molecular array and background-subtracted covariances for one or more pairs of data signal intensities;

- normalizing the data signal intensities, averages, and variances; and

- calculating ratios and variances of ratios of pairs of normalized data intensity signals.

59. (New) A system for automated extraction of data from a molecular array having features arranged in a regular pattern, the system comprising:

- a scanning component that produces images of the molecular array representing intensities of data signals emitted from discrete positions on a surface of the molecular array;

- a computer program that processes the images of the molecular array produced by the scanning component to index features in the images of the molecular array corresponding to molecules bound to features of the molecular array and that extracts data from the indexed features within images of the molecular array; and a computer for executing the computer program.

60. (New) A The system of claim 59 wherein data signal intensities emanating from discrete positions on the surface of the molecular array include:

radiation emitted by radioisotopes incorporated into molecules bound to features of the molecular array;

fluorescent emission from fluorophores incorporated into molecules bound to features of the molecular array; and

light emission from chemoluminescent moieties incorporated into molecules bound to features of the molecular array.

61. (New) The system of claim 59 wherein the computer program processes the images of the molecular array and extracts data from indexed features within images of the molecular array by:

receiving a number of images of the molecular array produced by the scanning component;

estimating initial positions of selected marker features within an image of the molecular array;

calculating refined positions of the selected marker features within the image of the molecular array;

using the refined positions of the selected marker features to compute an initial coordinate system for locating features of the molecular array in the number of images of the molecular array;

using the initial coordinate system to locate positions of strong features within one or more images of the molecular array;

refining the positions of strong features within the one or more images of the molecular array by analyzing data signal intensity values in regions of the one or more images of the molecular array that contain the strong features;

using the refined positions of strong features in the one or more images of the molecular array to calculate a refined coordinate system to locate positions of weak features within the number of images of the molecular array;

using the refined positions of strong features in the one or more images of the molecular array to calculate a refined coordinate system to locate positions of local background

regions surrounding all strong and weak features within the number of images of the molecular array; and

extracting data from strong features, and their respective local background regions, within the number of images of the molecular array using the refined positions of strong features within the number of images of the molecular array and extracting data from weak features, and their respective local background regions, within the number of images of the molecular array using locations for the weak features calculated from the refined coordinate system.

62. (New) The system of claim 61 wherein the computer program calculates background-subtracted averages, background-subtracted variances, and background-subtracted confidence intervals for data signal intensities integrated over features in the images corresponding to features of the molecular array.

63. (New) The system of claim 61 wherein the computer program calculates background-subtracted averages, background-subtracted variances, and background-subtracted confidence intervals for ratios of pairs of data signal intensities integrated over features in the images corresponding to features of the molecular array.